CLAIMS

Having thus described our invention, what we claim as new and desire to secure by Letters Patent is as follows:

1. A method for statistical regression using ensembles of classification solutions comprising the steps of:

running k-means clustering for k clusters on the set of values $\{y_i, i = 1 \ldots n\}$;

recording a mean value m_j of a cluster c_j for $j = 1 \ldots k$;

transforming regression data into classification data with a class label for an i-th case being a cluster number of y_i;

applying ensemble classifier and obtain a set of rules R; and

making a prediction for new case u, using a margin of M, where $0 \leq M \leq 1$.

2. The method recited in claim 1, wherein the step of making a prediction comprises the steps of:

applying all the rules R on the new case u;

for each class i, counting a number of satisfied rules (votes) v_i;

classifying t has the most votes, v_t;

considering a set of classes $P = \{p\}$ such that $v_p \geq M \cdot v_t$; and

generating a predicted output for case u, $y_u' = \dfrac{\sum_{j \in p} m_j v_j}{\sum_{j \in p} v_j}$.

3. A method of pattern recognition comprising the steps of:

applying clustering processes to determine a number of classes;

applying ensemble learning classification processes to predict most likely classes for a new example; and

then averaging regression values of most likely classes to predict a value of a new example.

4. A method of pattern recognition for a set of values, said method comprising the steps of:

determining a number of classes to be generated based on a trend of error of a class mean/median for the set of values;

classifying the values using ensemble learning classification and the determined number of classes;

generating a set of classification rules; and

averaging regression values of most likely classes to predict a value of a new example based on the set of rules.

5. A method of pattern recognition according to claim 4, wherein said step of determining a number of classes comprises the steps of:

determining the class mean/median for a variable number of classes;

determining a mean absolute deviation (MAD) based on the class means/medians; and

comparing the MAD to a predetermined percentage of MAD.

6. A method of pattern recognition according to claim 4, wherein the step of averaging regression values includes using margins for predicting the value of the new example.

7. A method of pattern recognition according to claim 4, wherein the step of averaging regression values comprises the steps of:

applying the set of classification rules to the new example;

for each class i, counting a number of satisfied rules (votes) v_i;

classifying t has the most votes, v_t;

considering a set of classes $P = \{p\}$ such that $v_p \geq M \cdot v_t$; and

generating a predicted output for case u, $y_u' = \frac{\sum_{j \in p} m_j v_j}{\sum_{j \in p} v_j}$.